Computing undetectable attacks on power grids

Daniel Bienstock and Mauro Escobar *

ABSTRACT

We consider combined data and physical attacks on power grids, motivated by recent events and research. We consider a setting where an attacker may alter the topology of a power grid by removing lines and may also alter the load (demand) of some nodes; simultaneously the attacker interdicts data flowing to the control center. We use the PMU model of data that provides high-fidelity AC power flow data (voltages and currents) The goal of the attacker is to provide data that paints a completely safe picture for the grid which is consistent with the net load change, while at the same time disguising large line overloads, a fundamentally dangerous situation that may lead to a cascading failure. We provide a computational procedure that efficiently computes sparse attacks even on cases of large grids.

1. INTRODUCTION

A question of current interest concerns the possibility that malicious attackers could interdict the data flow driving a power grid so as to cause planners to take incorrect actions; and whether this capability, combined with overt physical action damaging the grid could result in a catastrophic failure. Here we note that modern grids operate under a constant flow of data from sensors (some of which are located on nodes of the grid) to a centrally-located "control center" which under traditional grid operation, performs a computational procedure known as *state estimation* which computes an estimation of all real-time grid parameters on the basis of the available data, and uses the estimation to issue control commands to e.g. generators. Communication outside of this setup is nonstandard but may also be interdicted but is not considered in this paper.

This general topic has generated a huge amount of interest. Under the linearized, or DC model of power flows, [9], [7], [5] consider sparse data attacks and detection thereof. In [10], the authors consider a model where lines contained in a subgraph $\mathcal{H}$ known to the control center are removed while nodes in $\mathcal{H}$ are subjected to a data attack. Assuming a data-checking procedure that relies on the DC power flow model the authors discuss general conditions under which the physical attack and the correct data can be reconstructed (also see [11]). In [12], [13] the authors consider data under the nonlinear AC model, and again present general conditions under which the lines removed in $\mathcal{H}$ and the data attack can be efficiently identified.

In this paper we take the attacker's perspective and show that damaging attacks that remain *completely hidden*[1] can be quickly computed, even in the case of large grids. Specifically,

- The control center uses the AC power flow model to verify the consistency of data. We assume that a sensor reporting a (complex) voltage is located at each node of the network[2].
- In the physical component of the attack some lines may be removed and the load (demand) of some nodes may be altered. In the case of net load change, the process known as *secondary response* [6, 1, 2] will then cause the output of a subset of generator nodes (typically smaller generators) to react so as to alter their respective output in a pre-set proportion to the net load change; thus the attacker must mimic this behavior.
- The attacker will hijack a sparse subset of the sensors. The control center is unaware of the physical location of the attack. The data reported by the attacker is one where no line has been removed, and the implied power flow levels are all safe (each line's flow is within its limit). However, under the true system condition some lines experience large overloads.
- The size of the attack (number of affected nodes, in particular) is small relative to the size of the system.

We comment briefly on the above points. First, the requirement of consistency under the AC model is a stringent obstacle for the attacker, who will have to compute the solution to *two* AC power flow systems with additional constraints, a strongly NP-hard problem [3]. However, as pointed out in [7] the underlying computational expense may not be an obstacle to a sufficiently motivated attacker and, in fact, the assumption that the attacker is constrained to using polynomial-time algorithms may be a weakness on the

*Department of IEOR, Columbia University (email: dano,me2533@columbia.edu). Support from DOE, DTRA and DARPA is gratefully acknowledged.

[1]Note that [11] shows that under the DC model it is NP-Hard for the control center to identify an attack; a point here is that undetectability is a more stringent criterion that difficulty of computation.

[2]Below we outline the AC power flow model, but we stress that voltages suffice to describe the state of a known network.

part of the defending party. Indeed, in this paper we show that the computation can in fact be performed quite quickly even in the case of large systems with thousands of nodes and lines. At the same time, the AC consistency check also presents a challenge to the control center, because AC power flow systems of equations can have multiple solutions, unlike the case for DC power flows.

Second, the "attacker" in our analysis may indeed be fictitious; our results can also be interpreted to show that "natural" events that can affect both data sensors and grid hardware could give rise to difficult challenges to grid operators.

The rest of this abstract is organized as follows. In Section 2 we outline the mathematics of power flows.

2. POWER FLOWS

A grid, of transmission system, can be represented as an undirected graph where each line (edge) km has a complex *admittance* y_{km}. In operation of the grid each node k has a complex voltage (potential energy) $V_k = |V_k|e^{j\theta_k}$ where $j \doteq \sqrt{-1}$. For a line km the complex power from k to m equals the expression $S_{km} = V_k(y_{km}(V_k - V_m))^*$ which represents the laws of physics. The network equations satisfied by the voltages are of the form, for any node k,

$$\sum_{km\in\delta(k)} S_{km} \quad = \quad S_k^g - S_k^d. \tag{1}$$

In this expression $\delta(k)$ is the set of lines incident with node k, and S_k^g and S_k^d are the complex power generated at k and the load at k, respectively. Thus (1) states that the net balance of complex power at k is the difference between generation and demand at k. The *AC power flow* problem, in simplified form, seeks a solution to the equations (1), given generation and demands, and subject to the laws of physics and possibly additional constraints (e.g. bounds on each $|V_k|$ and upper bounds on each $|S_{km}|$). This computational task was proved to be strongly NP-hard in [3].

In part because of the computational difficulty arising in AC power flows, in operational practice it is common to use a linearized approximation. Here we deal with real power flows (i.e. ignoring the imaginary part of the S_{km} and of loads and generations). Instead of the (1) we obtain a system of linear equations of the form

$$B\theta \quad = \quad P^g - P^d \tag{2}$$

where the θ are the phase angles, as before, B is a generalized Laplacian matrix corresponding to the underlying network and using the inverse of reactances as weights (admittance y_{km} is the inverse of the *impedance* z_{km}, and reactance is the imaginary part of impedance). Denoting by x_{km} the reactance of km, the (real) power on km arising from a solution to (2) equals $(\theta_k - \theta_m)/x_{km}$ See e.g. [6, 1]. If the network is connected this power flow solution is unique, given vectors P^g and P^d and there is one degree of freedom in the θ.

Recent research on solution to AC power flow problems has identified a convex and frequently very close approximation [4], [8]. This is obtained by introducing, for each line km, new variables c_{km} and s_{km} representing, respectively, $|V_k||V_m|\cos(\theta_k - \theta_m)$ and $-|V_k||V_m|\sin(\theta_k - \theta_m)$, as well as a variable c_{kk} representing $|V_k|^2$ for each node k. Using these variables, the real part of (1) becomes

$$\sum_{km\in\delta(k)} [g_{km}c_{kk} - g_{km}c_{km}b_{km}s_{km}] \quad = \quad P_k^g - P_k^d. \tag{3}$$

where $y_{km} = g_{km} + jb_{km}$. A similar equation applies to the imaginary part of (1). Additionally one has

$$c_{km}^2 + s_{km}^2 \quad = \quad c_{kk}c_{km}, \quad \text{which is relaxed as} \tag{4}$$
$$c_{km}^2 + s_{km}^2 \quad \leq \quad c_{kk}c_{km}, \tag{5}$$

which is convex. The system of constraints (3) (plus its imaginary part correspondent), (5) as well as e.g. bounds on the c_{kk} used to represent bounds on $|V_k|^2$ yields a convex relaxation to AC power flows that in practice usually yields solutions with relatively small errors, i.e. that can be converted to full AC power flow solutions that are nearly feasible.

3. COMPUTING AN ATTACK

Let $\mathcal{N}$ denote the set of all nodes, $\mathcal{G}$ the set of generator nodes, and $\mathcal{R}$ the set of generators that participate in secondary response. This means that if a net load change of Δ takes place across the network, the output of generator $r \in \mathcal{R}$ will change by the amount $\alpha_r\Delta$, where $\alpha_r \geq 0$ is the pre-computed *participation factor* at r, and we have $\sum_{r\in\mathcal{R}} \alpha_r = 1$. Finally, we are given a set $\mathcal{A}$ of nodes whose data flow (voltage) and load can be altered by the attacker; we assume $\mathcal{A} \cap \mathcal{G} = \emptyset$. For brevity assume $\mathcal{A} = \mathcal{N} \setminus \mathcal{G}$ but more sophisticated choices are possible and will be described in the full paper.

Our attack computation proceeds as a three-step process.

First step. We first compute an attack under the DC power flow model. For a node i let P_i^g and P_i^d denote its initial, i.e. pre-attack active generation and load. For each line km we solve two linear programs, with constraints (6). In these constraints the $\breve{\theta}$ are reported to the control center whereas the $\bar{\theta}$ reflect the true state of the system. The variables in this formulation are all the $\breve{\theta}$, $\breve{P}^g$, $\breve{P}^d$ and $\hat{\theta}$, $\hat{P}^g$, $\hat{P}^d$.

$$B\breve{\theta} = \breve{P}^g - \breve{P}_d, \quad B\bar{\theta} = \bar{P}^g - \bar{P}_d \tag{6a}$$
$$\breve{P}_k^g \;=\; \bar{P}_k^g \;=\; P_k^g + \alpha_k \sum_i (\bar{P}_i^d - P_i^d) \quad \forall\, k \in \mathcal{R} \tag{6b}$$
$$\breve{P}_k^g \;=\; \bar{P}_k^g \;=\; P_k^g \quad \forall\, k \in \mathcal{G} \setminus \mathcal{R} \tag{6c}$$
$$\breve{P}_k^d \;=\; \bar{P}_k^d \;=\; P_k^d \quad \forall k \notin \mathcal{A} \tag{6d}$$
$$\breve{\theta}_k \;=\; \bar{\theta}_k \quad \forall k \notin \mathcal{A} \tag{6e}$$
$$|\breve{\theta}_k - \breve{\theta}_m|/x_{km} \;\leq\; \text{flow limit for } km,\ \forall \text{ line } km \tag{6f}$$
$$\text{all loads and generation nonnegative.} \tag{6g}$$

Here (6a) states the flow balance equation for the reported and true systems, (6b) states that generators participating in secondary response produce the correct (and same) output in both systems, (6c) states that the remaining generators do not change their output, (6d) and (6e) state that load nodes that are not attacked have unchanged load and phase angle in both systems (reported and true) and finally (6f) imposes limits on the power flows in the reported system. Now for each line km we solve two linear programs, both subject to (6); one maximizing and the other minimizing $\bar{\theta}_k - \bar{\theta}_m/x_{km}$. If either LP indicates a substantial violation of the line limit constraint for km then we have

computed an undetectable attack under the DC model and we move to the second step.

Second step. A successful attack computed in the first step will produce vectors $(\breve{P}^g, \breve{P}^d, \breve{\theta})$ and $(\bar{P}^g, \bar{P}^d, \bar{\theta})$ as well as a set $\mathcal{A}'$ of actually attacked nodes, i.e. the subset of $\mathcal{A}$ given by nodes i s.t. $(\breve{P}^g_i, \breve{P}^d_i, \breve{\theta}_i) \neq (\bar{P}^g_i, \bar{P}^d_i, \bar{\theta}_i)$. Typically, the set $\mathcal{A}'$ is sparse, i.e. rather small – this is a consequence of having solved a linear program.

We now attempt to *correct* the two vectors $(\breve{P}^g, \breve{P}^d, \breve{\theta})$ and $(\bar{P}^g, \bar{P}^d, \bar{\theta})$ into solutions for two separate AC power flow problems, one for the reported data case and one for the true data. In the reported case we *fix* the active generation at a bus k is set at $\breve{P}^g_k$, (plus or minus a small error), with reactive generation proportionally fixed at approximately $\breve{P}^g_k / P^g_k Q^g_k$, where Q^g_k is the initial (pre-attack) reactive the. We then attempt to solve these two power flow problems using Matpower [14]. A similar rule is used for the true data case. Either computation may fail (i.e. Matpower may fail to converge, a notorious problem). Assuming that both runs do converge we move to the third step.

Third step. Let $\breve{V}$ and $\bar{V}$ be the reported and true voltage vectors, respectively, which are the solutions to the two AC power flow problems solved in the second step. Even though we have stipulated that both AC problems have (nearly equal) generation at every generator node, and nearly equal loads at every node $k \notin \mathcal{A}'$ (implied by the constraints (6c)-(6e)) it may still be the case that there are small but non-trivial discrepancies in the voltages, and in particular the voltage phase angles

To correct these discrepancies we proceed as follows. Let $0 < \mu < 1$. Consider the voltage vector

$$\bar{V}^\mu \doteq \mu \breve{V} + (1-\mu)\bar{V}.$$

For $\mu \approx 1$ and any attacked bus k and line km we will have that

$$\begin{aligned} \bar{V}^\mu_k (y_{km}(\bar{V}^\mu_k - \bar{V}^\mu_m)^* &= \bar{V}_k (y_{km}(\bar{V}_k - \bar{V}_m)^* \\ &\quad + O(|y_{km}|(1-\mu)), \end{aligned} \tag{7}$$

since we always have $|V_k| \approx 1$ for every bus k. In other words, the power flow on any attacked line (or any line, for that matter) does not change much, proportionally, as we move from $\bar{V}$ to $\bar{V}^\mu$. And by construction of $\breve{V}$ and $\bar{V}$ we expect that the net complex power injection at each node does not change much. However, for $\mu < 1$ the discrepancy between $\breve{V}$ and $\bar{V}^\mu$ on non-attacked nodes will be smaller than that between $\breve{V}$ and $\bar{V}$.

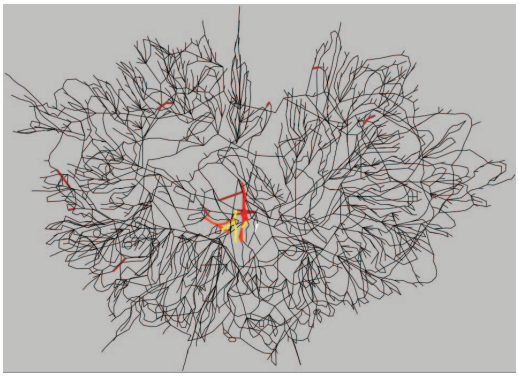
Accordingly, we construct the vector $\bar{V}^\mu$ for decreasing values of μ, and if we reach a vector with small enough discrepancy between $\breve{V}$ and $\bar{V}^\mu$ is small enough, and we still have substantial overloads, we have stop. At this point we have computed a successful attack.

Remark. As an alternative to this procedure, we can also consider an intermediate step where each of the two outputs (reported and true) of the pure-DC **Step 1** is first converted into a solution to the relaxation given by (3)-(5) which stipulates that non-attacked nodes have nearly equal values in the reported and true systems. The solution to the relaxation is then converted (as in **Steps 2** and **3** above) into a solution to the AC power flow problem. This process is smoother than the direct application of **Step 3** but may run into numerical difficulties, in that the solution of the convex relaxation (3)-(5) can prove challenging on large grids (even though convex!).

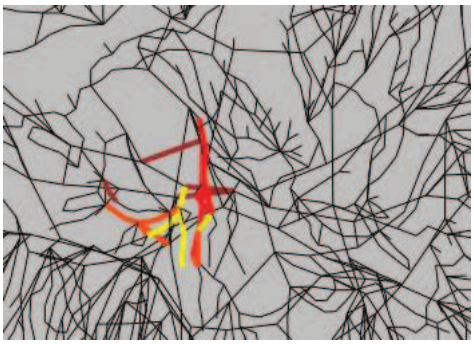
4. EXAMPLE

Here we outline an example using the "case2746wp" available from Matpower [14], with 2746 buses and 3514 branches. In this attack we trip one line, modify 168 loads in the neighborhood of this line and 89 bus data points.

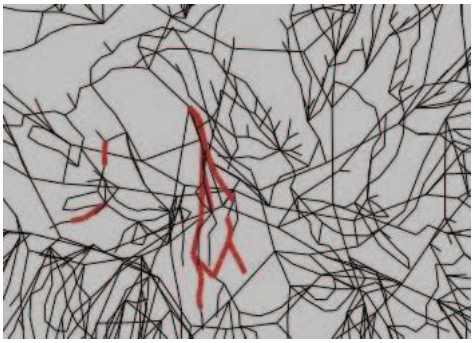
In the following figure, red and yellow colored branches indicate progressively higher actual line overloads (yellow = 100% overloads) while black lines indicate safely operating lines.



At the same time, the reported line flows are all safe (within line limits). The following zoomed-in figure shows the actual overloads:



For a comparison, the following figure shows the reported overloads:



Lines shown in bold are (at a maximum) at 80% load, and thus safe. Other lines are at lower load levels. In particular, the highest (true) overload lines are reported as being quite safe.

We will include more detailed analyses in examples in the full paper.

5. REFERENCES

[1] A. Bergen and V. Vittal. *Power Systems Analysis*. Prentice-Hall, 1999.

[2] D. Bienstock. *Electrical Transmission System Cascades and Vulnerability: An Operations Research Viewpoint*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2015.

[3] D. Bienstock and A. Verma. Strong NP-hardness of AC power flows feasibility. *arXiv:1512.07315*, 2015.

[4] C. Coffrin, H. L. Hijazi, and P. V. Hentenryck. The QC relaxation: Theoretical and computational results on optimal power flow. *CoRR*, abs/1502.07847, 2015.

[5] D. Deka, R. Baldick, and S. Vishwanath. Data attacks on power grids: Leveraging detection. In *2015 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5, Feb 2015.

[6] J. D. Glover, M. S. Sarma, and T. J. Overbye. *Power System Analysis and Design*. CENGAGE Learning, 2012.

[7] T. T. Kim and H. V. Poor. Strategic protection against data injection attacks on power grids. *IEEE Transactions on Smart Grid*, 2(2):326–333, June 2011.

[8] B. Kocuk, S. S. Dey, and X. A. Sun. Matrix Minor Reformulation and SOCP-based Spatial Branch-and-Cut Method for the AC Optimal Power Flow Problem. *CoRR*, abs/1703.03050, 2017.

[9] Y. Liu, P. Ning, and M. K. Reiter. False data injection attacks against state estimation in electric power grids. In *Proceedings of the 16th ACM Conference on Computer and Communications Security*, CCS '09, pages 21–32, New York, NY, USA, 2009. ACM.

[10] S. Soltan, M. Yannakakis, and G. Zussman. Power Grid State Estimation Following a Joint Cyber and Physical Attack. *IEEE Transactions on Control of Network Systems*, PP(99):1–1, 2016.

[11] S. Soltan, M. Yannakakis, and G. Zussman. REACT to Cyber Attacks on Power Grids. *Submitted*, 2017.

[12] S. Soltan and G. Zussman. EXPOSE the Line Failures following a Cyber-Physical Attack on the Power Grid. *Submitted*, 2017.

[13] S. Soltan and G. Zussman. Power Grid State Estimation after a cyber-physical attack under the AC power flow model. In *Proc. IEEE PES-GM'17*, 2017.

[14] R. D. Zimmerman, C. E. Murillo-Sánchez, and D. Gan. MATPOWER, A MATLAB Power System Simulation Package. *IEEE Trans. Power Sys.*, 26(1):12–19, 2011.